

基于 Bayes 判别法的马尾松毛虫一代、二代幼虫发生期的预报

钱广晶¹, 张书平¹, 宋学雨^{1,2}, 毕守东^{1*},
张国庆³, 邹运鼎², 方国飞⁴, 闫萍¹

(1. 安徽农业大学理学院, 合肥 230036; 2. 安徽农业大学林学与园林学院, 合肥 230036;
3. 安徽省潜山县林业局, 潜山 246300; 4. 国家林业局森林病虫害防治总站, 沈阳 110034)

摘要 为了提高马尾松毛虫 *Dendrolimus punctatus* (Walker) 发生量预测预报结果的准确性, 本文运用 Bayes 判别分析法建立安徽省潜山县 1983 年—2016 年 33 年的马尾松毛虫一代和二代幼虫发生期的预报模型。一代幼虫发生期的判别函数方程为: $f^{(1)} = -15\,744.058 - 361.501x_1 + 60.759x_2 + 133.502x_3 + 511.368x_4$; $f^{(2)} = -16\,854.938 - 375.596x_1 + 70.405x_2 + 132.608x_3 + 529.690x_4$; $f^{(3)} = -17\,645.295 - 384.956x_1 + 73.601x_2 + 134.955x_3 + 541.782x_4$; $f^{(4)} = -18\,179.639 - 382.408x_1 + 71.342x_2 + 135.234x_3 + 549.655x_4$ 对 1983 年—2018 年一代幼虫发生期预报结果历史符合率为 97.06%, 二代幼虫发生期的判别函数方程为: $f^{(1)} = -134\,898.483 + 559.235x_5 + 113.112x_6 - 250.033x_7 + 1\,461.350x_8$; $f^{(2)} = -138\,908.622 + 573.572x_5 + 118.340x_6 - 252.691x_7 + 1\,474.569x_8$; $f^{(3)} = -141\,430.680 + 577.358x_5 + 125.727x_6 - 254.610x_7 + 1\,483.336x_8$; $f^{(4)} = -143\,185.175 + 578.968x_5 + 129.628x_6 - 256.102x_7 + 1\,491.257x_8$ 对二代幼虫发生期的预报结果的历史符合率为 100%。对 2017 年和 2018 年的验证回报, 与实况结果一致。筛选出对预报量有密切关系的预报因子是本方法预报准确性的关键, 该方法是一种简便准确性高的预报方法。

关键词 马尾松毛虫幼虫; 发生期; Bayes 判别法; 预报

中图分类号: S 431 文献标识码: A DOI: 10.16688/j.zwbh.2019074

Forecast for the occurrence period of *Dendrolimus punctatus* larvae based on Bayes discriminant method for the first and second generations

QIAN Guangjing¹, ZHANG Shuping¹, SONG Xueyu^{1,2}, BI Shoudong^{1*},
ZHANG Guoqing³, ZOU Yunding², FANG Guofei⁴, YAN Ping¹

(1. School of Science, Anhui Agricultural University, Hefei 230036, China; 2. School of Forestry and Landscape Architecture, Anhui Agricultural University, Hefei 230036, China; Bureau of Forestry of Qianshan County, Anhui Province, Qianshan 246300, China; 4. Station for Prevention and Control of Forest Disease and Pests, Shenyang 110034, China)

Abstract To improve the accuracy of forecasting the occurrence of *Dendrolimus punctatus* Walker, the Bayes discriminant analysis method was used to predict the occurrence period of the first and second generations of *D. punctatus* larvae over a period of 33 years from 1983 to 2016 in Qianshan county, Anhui province. The discriminant function equation of the occurrence period of the first-generation larvae was as followed: $f^{(1)} = -15\,744.058 - 361.501x_1 + 60.759x_2 + 133.502x_3 + 511.368x_4$; $f^{(2)} = -16\,854.938 - 375.596x_1 + 70.405x_2 + 132.608x_3 + 529.690x_4$; $f^{(3)} = -17\,645.295 - 384.956x_1 + 73.601x_2 + 134.955x_3 + 541.782x_4$; $f^{(4)} = -18\,179.639 - 382.408x_1 + 71.342x_2 + 135.234x_3 + 549.655x_4$. The historical coincidence rate of the forecast results from 1983 to 2018 was 97.06%. The discriminant function equation for the second-generation larvae was as followed: $f^{(1)} = -134\,898.483 + 559.235x_5 + 113.112x_6 - 250.033x_7 + 1\,461.350x_8$; $f^{(2)} = -138\,908.622 + 573.572x_5 + 118.340x_6 - 252.691x_7 + 1\,474.569x_8$;

收稿日期: 2019-02-23 修订日期: 2019-04-27

基金项目: 国家林业公益性行业科研专项(201404410)

* 通信作者 E-mail: bishoudong@163.com

$f^{(3)} = -141\,430.680 + 577.358x_5 + 125.727x_6 - 254.610x_7 + 1\,483.336x_8$; $f^{(4)} = -143\,185.175 + 578.968x_5 + 129.628x_6 - 256.102x_7 + 1\,491.257x_8$. The historical coincidence rate of the forecast results for the second-generation larvae from 1983 to 2018 was 100%. The verification returns for 2017 and 2018 were consistent with the observed data. Screening out the forecasting factors closely related to the forecasting quantity was the key to the accuracy of forecast. This forecasting method is simple and accurate.

Key words *Dendrolimus punctatus* larvae; period of occurrence; the Bayes discriminant analysis; forecast

马尾松毛虫 *Dendrolimus punctatus* (Walker) 分布于中国皖、豫、川、黔、陕、滇、赣、湘、浙、闽、粤、琼、桂等省(区),主要为害马尾松 *Pinus massoniana* Lamb., 还为害黑松 *P. thunbergii* Parl.、火炬松 *P. taeda* L.、湿地松 *P. elliotii* Engelm.、晚松 *P. rigida* var. *serotina* (Michx.) Loud ex Hoopes、海南五针松 *P. fenzeliana* Hand.-Mazz. 等松属植物。20 世纪中叶在中国森林害虫中马尾松毛虫是发生最广、危害面积最大、经常猖獗成灾的害虫。在广大丘陵地区虫害此起彼伏,针叶常被吃光,被害时如同火烧,造成了巨大的经济效益损失和生态效益损失。再者,人们在从事林业活动中接触马尾松毛虫毒毛,容易引发皮炎和关节肿痛,该虫不但影响林业生产,还危害人类健康^[1-4]。进入 21 世纪,由于各地采用封山育林、混交、间作等措施优化了森林生态环境,科学地运用综合治理措施,使马尾松毛虫的危害得到有效的控制,但该虫具有强大的繁殖潜力,遇到有利的生态环境极易暴发成灾,对其的监测不能放松警惕。

马尾松毛虫在中国一年发生 2~4 代,发生世代的多少,随不同地方而异,在河南省信阳地区一年发生 2 代为主,在长江流域诸省一年发生 2~3 代,而在广东、广西、福建南部一年发生 3~4 代,海南一年发生 4~5 代^[3]。安徽潜山县一年发生 3 代,即 4~6 月上旬为越冬代,6 月上旬~8 月中下旬为一代,8 月中下旬~12 月为第二代。马尾松毛虫发生的预测预报是对其进行综合防治的基本工作。科研工作者^[5-12]分别采用不同方法预测马尾松毛虫的发生量、虫害等级、发生类别、发生空间格局,为马尾松毛虫的综合防治工作提供了有力的支持。由于各地气象条件、植被条件和地形地貌等不同,马尾松毛虫的发生特点也不完全相同。马尾松毛虫幼虫一个世代的累计发生量直接与该虫的防治有关,为了有效防治马尾松毛虫,本研究采用 Bayes 判别分析法研究马尾松毛虫一、二代幼虫的发生期,以期马尾松毛虫的综合治理提供科学依据。

1 材料与方法

1.1 材料来源

马尾松毛虫资料来自中国国家林业局与森林病虫害防治总站,气象资料来自国家气象局,资料的时间跨度为 1983 年—2016 年。根据国家林业局森林病虫害防治总站制定的调查方法进行调查,即采用踏查和详查相结合的办法,对不同虫态采取不同的观测方法。沿林班线、林道、公路、铁路等线路调查,目测发生范围,为害状况,发生虫情或灾情立即设临时标准地,采取平行线抽样法抽取 20 株标准株详查。卵期调查是在雌蛾羽化高峰后 1~2 d 调查平均卵块数,卵块平均粒数;幼虫期调查,1~2 龄幼虫调查枯黄卷曲的枝数,推算幼虫数,3 龄以上幼虫 3 m 以下小树直接调查合计树冠上的幼虫数,大树用“虫粪粒推算法”调查,幼虫越冬期调查树干基部树皮缝中的幼虫数推算全部虫口。蛹期调查,在结茧盛期后 2~5 d 剖茧,调查雌雄性比、平均雌蛹重、死亡率、寄生率;成虫期调查,在成虫羽化前 2~3 d 及羽化期用黑光灯诱集,统计其诱集数量和雌雄性比。

1.2 马尾松毛虫的发生期(y)的分级标准

一代:1 级, $66 \text{ d} \leq y \leq 67 \text{ d}$; 2 级, $68 \text{ d} \leq y \leq 69 \text{ d}$; 3 级, $70 \text{ d} \leq y \leq 71 \text{ d}$; 4 级, $72 \text{ d} \leq y \leq 73 \text{ d}$;

二代:1 级, $158 \text{ d} \leq y \leq 159 \text{ d}$; 2 级, $160 \text{ d} \leq y \leq 161 \text{ d}$; 3 级, $162 \text{ d} \leq y \leq 163 \text{ d}$; 4 级, $164 \text{ d} \leq y \leq 165 \text{ d}$ 。

1.3 Bayes 判别分析预报法

判别分析又叫统计分析法,它可以利用多个因子对预报对象作出客观的分类预报,常用的是利用 Fisher 准则建立判别函数进行预测,本文用 Bayes 准则作出预报^[13-15], Bayes 准则的基本思想是,当预报量分为 G 类($G > 2$),则相应的预报因子按预报量的类别也分为 G 组,对应于预报量第 g 类($g=1, 2, \dots, G$),由 m 个预报因子 N_g 次观测值组成的样本为:

$$\begin{array}{cccc} x_{11}^{(g)} & x_{21}^{(g)} & \cdots & x_{m1}^{(g)} \\ x_{12}^{(g)} & x_{22}^{(g)} & \cdots & x_{m2}^{(g)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{1N_g}^{(g)} & x_{2N_g}^{(g)} & \cdots & x_{mN_g}^{(g)} \end{array}$$

其中 $x_{ki}^{(g)}$ 为对应 y 为 g 类,第 i 个样本($i=1, 2, \dots, N_g$)中预报因子 x_k ($k=1, 2, \dots, m$) ($k=1, 2, \dots, m$) 的值。设 $\bar{x}_k^{(g)}$ 为第 g 类样本因子 x_k 的平均数,即:

$$\bar{x}_k^{(g)} = \frac{1}{N_g} \sum_{i=1}^{N_g} x_{ki}^{(g)} \quad (g=1, 2, \dots, G; k=1, 2, \dots, m) \quad (1)$$

这些平均值组成 G 个 m 向量

$$\begin{aligned} \bar{x}^{(1)} &= (\bar{x}_1^{(1)}, \bar{x}_2^{(1)}, \dots, \bar{x}_m^{(1)}) \\ \bar{x}^{(2)} &= (\bar{x}_1^{(2)}, \bar{x}_2^{(2)}, \dots, \bar{x}_m^{(2)}) \\ &\cdots \quad \cdots \quad \cdots \quad \cdots \\ \bar{x}^{(i)} &= (\bar{x}_1^{(i)}, \bar{x}_2^{(i)}, \dots, \bar{x}_m^{(i)}) \end{aligned} \quad (2)$$

它们的列向量为:

$$\begin{aligned} \bar{x}^{(1)'} &= \begin{bmatrix} \bar{x}_1^{(1)} \\ \bar{x}_2^{(1)} \\ \vdots \\ \bar{x}_m^{(1)} \end{bmatrix}, \bar{x}^{(2)'} = \begin{bmatrix} \bar{x}_1^{(2)} \\ \bar{x}_2^{(2)} \\ \vdots \\ \bar{x}_m^{(2)} \end{bmatrix}, \dots, \\ \bar{x}^{(G)'} &= \begin{bmatrix} \bar{x}_1^{(G)} \\ \bar{x}_2^{(G)} \\ \vdots \\ \bar{x}_m^{(G)} \end{bmatrix} \end{aligned} \quad (3)$$

在作预报时,预报因子的数值可用向量 x 表示。

$$x = (x_1, x_2, \dots, x_m) \quad (4)$$

对于(4)所示的预报因子向量 x ,如何判断它属于何类时预示 y 属于何类。实质上就是 m 维空间进行划分,分为 G 个部分,使每一部分都与预报量某类相对应。在划分中由于错分而引起预报不准,在使平均错分损失最小的原则下,寻找一种最佳的划分,这就是 Bayes 准则。具体地说,如果 P_g 表示 y 为第 g 类的发生概率, $f_g(x)$ 为对应于第 g 类的概率密度。Bayes 准则可表达为 $P_g f_g(x)$ 取最大值。因此,是求出类似于各类出现概率那样的量。其构成的判别函数为:

$$f^{(g)} = c_0^{(g)} + c_1^{(g)} x_1 + c_2^{(g)} x_2 + \cdots + c_m^{(g)} x_m \quad (g=1, 2, \dots, G) \quad (5)$$

以 $f^{(g)}$ 的数值来反映 y 出现在第 g 类可能性的。大小。 $f^{(g)}$ 值越大,表示 y 出现 g 类的可能性越大。

其判别系数向量形式为:

$$(c_1^{(g)} \ c_2^{(g)} \cdots c_m^{(g)}) = (N-G) s^{-1} \bar{x}^{(g)'} \quad (6)$$

$$c_0^{(g)} = \ln \frac{N_g}{N} - \frac{1}{2} \sum_{k=1}^m c_k^{(g)} \bar{x}_k^{(g)} \quad (g=1, 2, \dots, m) \quad (7)$$

其中 N —总样本数; N_g —第 g 类的样本数; G —所分级数;

$\bar{x}^{(g)'} - \bar{x}^{(g)}$ 的向量。即:

$$\bar{x}^{(g)'} = \begin{bmatrix} \bar{x}_1^{(g)} \\ \bar{x}_2^{(g)} \\ \vdots \\ \bar{x}_m^{(g)} \end{bmatrix} \quad (8)$$

s^{-1} 为因子间关联数矩阵 s 的逆矩阵

$$s = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1m} \\ s_{21} & s_{22} & \cdots & s_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ s_{m1} & s_{m2} & \cdots & s_{mm} \end{bmatrix} \quad (9)$$

(10)式中 s_{kl} 便成为 x_k 与 x_l 之间的关联数。

$$s_{kl} = \sum_{g=1}^G s_{kl}^g = s_{kl}^{(1)} + s_{kl}^{(2)} + \cdots + s_{kl}^{(G)} = \sum_{g=1}^G \sum_{i=1}^{N_g} (x_{ki}^{(g)} - \bar{x}_k^{(g)})(x_{li}^{(g)} - \bar{x}_l^{(g)}) \quad (10)$$

对于第 g 类的样本,则有

$$s_{kl}^{(g)} = s_{kl}^{(g)} = \sum_{i=1}^{N_g} (x_{ki}^{(g)} - \bar{x}_k^{(g)})(x_{li}^{(g)} - \bar{x}_l^{(g)}) \quad (11)$$

本文所有运算用 SPSS 程序进行计算。

2 结果与分析

根据安徽省潜山县马尾松毛虫 1983 年—2016 年的资料统计,一年发生 3 代,即越冬代、一代和二代。1—6 月为越冬代,1—4 月下旬为幼虫期,4 月下旬—5 月中下旬为蛹期,5 月下旬—6 月上旬为成虫期,5 月中下旬—8 月中旬为第一代,5 月中下旬—6 月上旬为卵期,6 月上旬—7 月中下旬为幼虫期,7 月中下旬—8 月上旬为蛹期,8 月上旬—8 月中旬为成虫期,8 月中下旬—12 月为第二代,8 月中下旬—9 月上旬为卵期,9 月上旬—11 月上旬为幼虫期,温度适宜时少量幼虫可继续发育。

2.1 马尾松毛虫一代幼虫发生期的预报

将安徽省潜山县 1983 年—2016 年 34 年的马尾松毛虫一代幼虫发生期及其预报值列于表 1。

表 1 一代幼虫高峰期 y_1 及自变量因子¹⁾

Table 1 Peak period y_1 and independent variable factors for first-generation larvae

年份 Year	一代幼虫高峰期 Peak period of the first-generation larvae		越冬代蛹高峰期 Peak period of overwintering pupa		越冬代成虫高峰期 Peak period of overwintering adult		一代卵始盛期 Initial peak period for the first- generation egg		一代幼虫始盛期 Initial peak period of the first-generation larvae		分级 Grade	回报值 Return value	评判 Judgement
	日期/ 月-日 Date	y	日期/ 月-日 Date	x_1	日期/ 月-日 Date	x_2	日期/ 月-日 Date	x_3	日期/ 月-日 Date	x_4			
1983	06-12	73	05-09	39	05-23	53	06-03	64	06-10	71	4	4	✓
1984	06-12	73	05-09	39	05-23	53	06-03	64	06-10	71	4	4	✓
1985	06-12	73	05-09	39	05-23	53	06-03	64	06-10	71	4	4	✓
1986	06-12	73	05-08	38	05-23	53	06-03	64	06-09	70	4	4	✓
1987	06-12	73	05-08	38	05-23	53	06-03	64	06-08	69	4	4	✓
1988	06-12	73	05-07	37	05-24	54	06-03	64	06-08	69	4	4	×
1989	06-10	71	05-06	36	05-24	54	06-03	64	06-07	68	3	3	✓
1990	06-10	71	05-06	36	05-23	53	06-01	62	06-07	68	3	3	✓
1991	06-10	71	05-06	36	05-23	53	06-03	64	06-07	68	3	3	✓
1992	06-09	70	05-06	36	05-23	53	06-02	63	06-06	67	3	3	✓
1993	06-09	70	05-05	35	05-23	53	06-02	63	06-06	67	3	3	✓
1994	06-08	69	05-05	35	05-22	52	06-01	62	06-05	66	2	2	✓
1995	06-08	69	05-04	34	05-21	51	05-30	60	06-05	66	2	2	✓
1996	06-08	69	05-03	33	05-18	48	05-29	59	06-05	66	2	2	✓
1997	06-08	69	05-02	32	05-15	45	05-27	57	06-05	66	2	2	✓
1999	06-08	69	05-02	32	05-15	45	05-27	57	06-05	66	2	2	✓
2000	06-06	67	05-01	31	05-12	42	05-26	56	06-03	64	1	1	✓
2001	06-06	67	04-30	30	05-09	39	05-25	55	06-03	64	1	1	✓
2002	06-06	67	04-30	30	05-09	39	05-25	55	06-03	64	1	1	✓
2003	06-06	67	04-29	29	05-08	38	05-25	55	06-02	63	1	1	✓
2004	06-06	67	04-29	29	05-08	38	05-25	55	06-02	63	1	1	✓
2005	06-06	67	04-29	29	05-08	38	05-25	55	06-02	63	1	1	✓
2006	06-06	67	04-27	27	05-06	36	05-23	53	06-02	63	1	1	✓
2007	06-05	66	04-27	27	05-04	34	05-22	52	06-02	63	1	1	✓
2008	06-05	66	04-27	27	05-04	34	05-20	50	06-02	63	1	1	✓
2009	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2010	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2011	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2012	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2013	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2014	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2015	06-05	66	04-25	25	05-02	32	05-18	48	06-02	63	1	1	✓
2016	06-05	66	04-25	25	05-02	32	05-16	46	06-02	63	1	1	✓

1) 一代幼虫高峰期分级,6月6日以前,1级;6月7日-8日,2级;6月9日-10日,3级;6月11日,4级。
The level standards of peak period for the first-generation larvae are as follows: Level 1 occurs before June 6; level 2 occurs from June 7 to June 8; level 3 occurs from June 9 to June 10, and level 4 occurs at June 11.

经过筛选,越冬代蛹高峰期(x_1)、越冬代成虫高峰期(x_2)、一代卵始期(x_3)、一代幼虫始盛期(x_4)与一代幼虫高峰期(y_1)之间的相关系数 r 依次为 0.960 2, 0.925 1, 0.920 6 和 0.983 7, $df=33$ 时, $r_{0.05}=0.330$, $r_{0.01}=0.40$, r 值均大于 $r_{0.01}$ (0.436 6)。将 x_1-x_4 的值也列于表 1。按照(1)式求得自变量平均值的列向量为:

$$\bar{x}_1^{(1)} = \begin{bmatrix} 27.000\ 0 \\ 34.941\ 2 \\ 51.058\ 8 \\ 63.176\ 5 \end{bmatrix}, \bar{x}_2^{(2)} = \begin{bmatrix} 33.200\ 0 \\ 48.200\ 0 \\ 59.000\ 0 \\ 66.000\ 0 \end{bmatrix},$$

$$\bar{x}_3^{(3)} = \begin{bmatrix} 35.800 & 0 \\ 53.200 & 0 \\ 63.200 & 0 \\ 67.600 & 0 \end{bmatrix}, \bar{x}_4^{(4)} = \begin{bmatrix} 38.333 & 3 \\ 53.166 & 7 \\ 64.000 & 0 \\ 70.166 & 7 \end{bmatrix}$$

根据公式(10)和(11)求得 x_1 、 x_2 、 x_3 、 x_4 的关联数矩阵:

$$s = \begin{bmatrix} 88.933 & 3 & 133.666 & 7 & 131.200 & 0 & 14.266 & 7 \\ 133.666 & 7 & 227.374 & 5 & 210.858 & 8 & 14.409 & 8 \\ 131.200 & 0 & 210.858 & 8 & 219.741 & 2 & 13.223 & 5 \\ 14.266 & 7 & 14.409 & 8 & 13.223 & 5 & 8.503 & 9 \end{bmatrix}$$

求出上述关联数矩阵 s 的逆矩阵:

$$s^{-1} = \begin{bmatrix} 0.270 & 5 & -0.069 & 6 & -0.082 & 1 & 0.208 & 0 \\ -0.069 & 6 & 0.058 & 5 & -0.017 & 2 & 0.044 & 5 \\ -0.082 & 1 & -0.017 & 2 & 0.066 & 3 & -0.063 & 9 \\ -0.208 & 0 & 0.044 & 5 & -0.063 & 9 & 0.291 & 7 \end{bmatrix}$$

根据公式(6)求出:

$$\begin{aligned} (c_1^{(1)}, c_2^{(1)}, c_3^{(1)}, c_4^{(1)}) &= \\ (-361.501, 60.759, 133.502, 511.368) \\ (c_1^{(2)}, c_2^{(2)}, c_3^{(2)}, c_4^{(2)}) &= \\ (-375.596, 70.405, 132.608, 529.690) \\ (c_1^{(3)}, c_2^{(3)}, c_3^{(3)}, c_4^{(3)}) &= \\ (-384.956, 73.601, 134.955, 541.782) \\ (c_1^{(4)}, c_2^{(4)}, c_3^{(4)}, c_4^{(4)}) &= \\ (-382.408, 71.342, 135.234, 549.655) \end{aligned}$$

根据公式(7)求得:

$$\begin{aligned} c_0^{(1)} &= -15\,744.058, c_0^{(2)} = -16\,854.938, \\ c_0^{(3)} &= -17\,645.295, c_0^{(4)} = -18\,179.639. \end{aligned}$$

根据公式(5)列出判别函数方程:

$$\begin{aligned} f^{(1)} &= -15\,744.058 - 361.501x_1 + \\ &60.759x_2 + 133.502x_3 + 511.368x_4; \\ f^{(2)} &= -16\,854.938 - 375.596x_1 + \\ &70.405x_2 + 132.608x_3 + 529.690x_4; \\ f^{(3)} &= -17\,645.295 - 384.956x_1 + \\ &73.601x_2 + 134.955x_3 + 541.782x_4; \\ f^{(4)} &= -18\,179.639 - 382.408x_1 + \\ &71.342x_2 + 135.234x_3 + 549.655x_4. \end{aligned}$$

根据 1983 年—2016 年安徽省潜山县与马尾松毛虫一代幼虫高峰期关系极相关的越冬代蛹高峰期(x_1)、越冬代成虫高峰期(x_2)、一代卵始期(x_3)和一代幼虫始盛期(x_4)代入上述判别函数方程。如 1983 年, x_1 为 39, x_2 为 53, x_3 为 64, x_4 为 71, 代入判别函数方程 $f^{(1)} = 18\,228.886$, $f^{(2)} = 18\,323.185$, $f^{(3)} = 18\,345.916$, $f^{(4)} = 18\,368.056$; 根据 Bayes 准则可表达为 $P_g f_g^{(x)}$ 取最大值, $f^{(4)}$ 值最大, 即 1983 年马尾松毛虫一代幼虫发生期为 4 级。对 34 年的预报结果进行回报检验, 结果也列于表 1。并对 2017 年和 2018 年的预报结果进行检验, 1983 年—2018 年的历史符合率为 97.06%。

2.2 二代幼虫发生期的预报

将安徽省潜山县 1983 年—2016 年 34 年的马尾松毛虫二代幼虫发生期及其预报值列于表 2。

表 2 二代幼虫高峰期 y_2 及自变量因子¹⁾

Table 2 Peak occurrence period y_2 and independent variable factors for second-generation larvae

年份 Year	二代幼虫高峰期 Peak period for the second- generation larvae		一代蛹高峰期 Peak period for the first- generation pupa		一代成虫高峰期 Peak period for the first- generation adult		二代卵高峰期 Peak period for the second- generation egg		二代幼虫始盛期 Initial peak period for the second- generation larvae		分级 Grade	回报值 Return value	评判 Judgement
	日期/ 月-日		日期/ 月-日		日期/ 月-日		日期/ 月-日		日期/ 月-日				
	Date	y	Date	x ₁	Date	x ₂	Date	x ₃	Date	x ₄			
1983	09-11	164	07-27	118	08-13	135	08-28	150	09-07	160	4	4	✓
1984	09-11	164	07-27	118	08-13	135	08-27	149	09-07	160	4	4	✓
1985	09-11	164	07-26	117	08-13	135	08-27	149	09-07	160	4	4	✓
1986	09-11	164	07-28	119	08-13	135	08-27	149	09-07	160	4	4	✓
1987	09-11	164	07-28	119	08-13	135	08-27	149	09-07	160	4	4	✓
1988	09-11	164	07-28	119	08-13	135	08-28	150	09-07	160	4	4	✓
1989	09-10	163	07-27	118	08-12	134	08-28	150	09-06	159	3	3	✓
1990	09-10	163	07-27	118	08-12	134	08-26	148	09-06	159	3	3	✓
1991	09-10	163	07-27	118	08-11	133	08-26	148	09-06	159	3	3	✓
1992	09-10	163	07-27	118	08-10	132	08-27	149	09-06	159	3	3	✓
1993	09-08	161	07-26	117	08-10	132	08-27	149	09-05	158	2	2	✓
1994	09-07	160	07-26	117	08-09	131	08-26	148	09-05	158	2	2	✓

续表 2 Table 2(Continued)

年份 Year	二代幼虫高峰期 Peak period for the second- generation larvae		一代蛹高峰期 Peak period for the first- generation pupa		一代成虫高峰期 Peak period for the first- generation adult		二代卵高峰期 Peak period for the second- generation egg		二代幼虫始盛期 Initial peak period for the second- generation larvae		分级 Grade	回报率 Return value	评判 Judgement
	日期/ 月-日		日期/ 月-日		日期/ 月-日		日期/ 月-日		日期/ 月-日				
	Date	y	Date	x_1	Date	x_2	Date	x_3	Date	x_4			
1995	09-07	160	07-26	117	08-08	130	08-25	147	09-05	158	2	2	✓
1996	09-07	160	07-26	117	08-08	130	08-25	147	09-05	158	2	2	✓
1997	09-07	160	07-26	117	08-07	129	08-24	146	09-04	157	2	2	✓
1999	09-07	160	07-26	117	08-07	129	08-24	146	09-04	157	2	2	✓
2000	09-06	159	07-24	115	08-06	128	08-22	144	09-03	156	1	1	✓
2001	09-06	159	07-23	114	08-05	127	08-24	146	09-03	156	1	1	✓
2002	09-06	159	07-23	114	08-05	127	08-24	146	09-03	156	1	1	✓
2003	09-06	159	07-22	113	08-05	127	08-21	143	09-03	156	1	1	✓
2004	09-06	159	07-22	113	08-05	127	08-21	143	09-03	156	1	1	✓
2005	09-06	159	07-22	113	08-05	127	08-21	143	09-03	156	1	1	✓
2006	09-06	159	07-22	113	08-05	127	08-21	143	09-03	156	1	1	✓
2007	09-06	159	07-22	113	08-05	127	08-21	143	09-03	156	1	1	✓
2008	09-06	159	07-22	113	08-05	127	08-19	141	09-03	156	1	1	✓
2009	09-06	159	07-22	113	08-05	127	08-17	139	09-03	156	1	1	✓
2010	09-06	159	07-22	113	08-05	127	08-17	139	09-03	156	1	1	✓
2011	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓
2012	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓
2013	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓
2014	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓
2015	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓
2016	09-06	159	07-22	113	08-05	127	08-17	139	09-02	155	1	1	✓

1) 二代幼虫高峰期分级,9月6日以前,1级;9月7日-8日,2级;9月9日-10日,3级;9月11日,4级。

The level standards of peak period for the second-generation larvae are as follows: Level one occurs before September 6; level 2 occurs from September 7 to September 8; level 3 occurs from September 9 to September 10, and level 4 occurs at September 11.

经过筛选一代累计虫口(x_5)、二代卵盛期发生量(x_6)、二代幼虫高峰期发生量(x_7)、二代卵量(x_8)和二代幼虫高峰期(y_2)的相关系数依次为 0.754 0、0.960 4、0.990 2 和 0.999 9, r 值均大于 $r_{0.01}$ (0.430), x_5-x_8 与二代幼虫发生期之间均极相关。将 x_5-x_8 的值也列于表 2。

按(1)式求出自变量平均值的列向量为:

$$x_5^{(1)} = \begin{bmatrix} 113.235\ 3 \\ 127.058\ 8 \\ 141.411\ 8 \\ 155.647\ 1 \end{bmatrix}, x_6^{(2)} = \begin{bmatrix} 117.000\ 0 \\ 130.166\ 7 \\ 147.166\ 7 \\ 157.666\ 7 \end{bmatrix},$$
$$x_7^{(3)} = \begin{bmatrix} 118.000\ 0 \\ 133.250\ 0 \\ 148.750\ 0 \\ 159.000\ 0 \end{bmatrix}, x_8^{(4)} = \begin{bmatrix} 118.333\ 3 \\ 135.000\ 0 \\ 149.333\ 3 \\ 160.000\ 0 \end{bmatrix}$$

根据公式(10)和(11)求出 x_5, x_6, x_7, x_8 的关联数矩阵:

$$s = \begin{bmatrix} 8.392\ 2 & 1.764\ 7 & 14.686\ 3 & 1.411\ 8 \\ 1.764\ 7 & 10.524\ 5 & 9.671\ 6 & 2.686\ 3 \\ 14.686\ 3 & 9.671\ 6 & 119.034\ 3 & 16.803\ 9 \\ 1.411\ 8 & 2.686\ 3 & 16.803\ 9 & 5.215\ 7 \end{bmatrix}$$

求出上述关联数矩阵的逆矩阵:

$$s^{-1} = \begin{bmatrix} 0.157\ 7 & -0.015\ 2 & -0.024\ 4 & 0.043\ 7 \\ -0.015\ 2 & 0.111\ 1 & 0.000\ 6 & -0.055\ 1 \\ -0.024\ 4 & 0.000\ 6 & 0.019\ 2 & -0.055\ 6 \\ 0.043\ 7 & -0.055\ 1 & -0.055\ 6 & 0.387\ 5 \end{bmatrix}$$

根据公式(6)求出 $c_5^{(g)}, c_6^{(g)}, c_7^{(g)}, c_8^{(g)}$, 根据公式(7)求出 $c_0^{(1)}, c_0^{(2)}, c_0^{(3)}, c_0^{(4)}$ 依次为 -134 898.483, -138 908.622, -141 430.680 和 -143 185.175, 根据公式(5)列出判别函数方程:

$$f^{(1)} = -134\ 898.483 + 559.235x_5 + 113.112x_6 - 250.033x_7 + 1461.350x_8;$$
$$f^{(2)} = -138\ 908.622 + 573.572x_5 + 118.340x_6 - 252.691x_7 + 1\ 474.569x_8;$$
$$f^{(3)} = -141\ 430.680 + 577.358x_5 +$$

$$125.727x_6 - 254.610x_7 + 1483.336x_8;$$

$$f^{(4)} = -143185.175 + 578.968x_5 +$$

$$129.628x_6 - 256.102x_7 + 1491.257x_8。$$

将 1983 年—2016 年安徽省潜山县与马尾松毛虫二代幼虫高峰期极相关的一代蛹高峰期(x_5)、一代成虫高峰期(x_6)、二代卵高峰期(x_7)和二代幼虫始盛期(x_8)代入上述判别函数方程,如 1983 年 $x_5=118$, $x_6=135$, $x_7=150$, $x_8=160$,求 $f^{(g)}$ 值。 $f^{(1)}=142672.417$, $f^{(2)}=142776.164$, $f^{(3)}=142812.969$, $f^{(4)}=142818.649$,根据 Bayes 准则可表达为 $P_g f_g^{(x)}$ 取最大值, $f^{(4)}$ 值最大。即 1983 年马尾松毛虫二代幼虫发生期为 4 级,对 34 年的预报结果进行回报检验,计算结果也列于表 2 并对 2018 年的预报进行检验,即回报 2018 年马尾松毛虫二代幼虫发生期为 1 级,实况也是 1 级,则 1983 年—2018 年的历史符合率为 100%。

3 结论与讨论

用 Bayes 判别法的判别函数方程预报安徽省潜山县一、二代马尾松毛虫幼虫发生期,结果表明,1983 年—2018 年预报马尾松毛虫一代幼虫发生期的历史符合率为 97.06%;1983 年—2018 年预报马尾松毛虫二代幼虫发生期的历史符合率为 100%。

对马尾松毛虫一、二代幼虫发生期预报结果准确与否及历史符合率的高低,一是取决于对预报量的分级是否科学,本文将马尾松毛虫一、二代发生期分为 4 级,相邻级之间级差为 1 d,这是综合了安徽省潜山县 35 年的发生期情况综合分析制定的,再者筛选自变量因子,影响马尾松毛虫幼虫发生期有很多因子,诸如气候因子、天敌等生物因子以及马尾松毛虫各虫态之间的关系,判别分析法也采用逐步回归法筛选因子。本文通过单相关分析各因子与马尾松毛虫幼虫发生期相关性大小。筛选的 x_1-x_8 都是与马尾松毛虫幼虫发生期达到极相关的因子,因此预报结果较为理想。预报量之一即一代幼虫高峰期(y_1)与预报因子越冬代蛹高峰期(x_1)、越冬代成虫高峰期(x_2)、一代卵始盛期(x_3)和一代幼虫始盛期(x_4)以及另一预报量二代幼虫高峰期(y_2)与预报因子一代蛹高峰期(x_5)、一代成虫高峰期(x_6)、二代卵高峰期(x_7)和二代幼虫始盛期(x_8)都是不同虫态发生期之间的关系,发生期的早迟却与当时的温度有关,马尾松毛虫和其他变温动物一样,在其生长发育过程中必须要有一定数量的有效积温,这个有效

积温基本上是一个常数,由于在该虫生长发育过程中温度会有起伏变化,因此一些预报因子如越冬代蛹高峰期等,若在此时段温度显著低于常年,那么越冬代蛹高峰期就要推迟,若显著高于常年,则越冬代蛹高峰期就会提前,其他虫态发生期也会出现同样的情况,马尾松毛虫的生长发育是由卵、幼虫、蛹、成虫的顺序变化的,前面虫态发生期的变化势必影响下一虫态发生期的变化,由于预报因子推迟(或提前),那么预报量也相应地推迟(或提前)。

若预报的越冬代蛹高峰期(或其他虫态发生期)时段温度在常年变化幅度内,越冬代蛹高峰期及下面的虫态发生期与常年相比,变化幅度很小,这时预报结果一代幼虫高峰期与实况差异很小,若预报的越冬代蛹高峰期温度异常,其预报结果与实况可能误差很大。二代幼虫高峰期的情况也是如此。

周夏芝等^[16]曾用回归预测法、平稳时间序列法、BP 神经网络法、马尔科夫链法和列联表多因子分级综合相关法预报安徽潜山县马尾松毛虫一、二代幼虫高峰期发生量,其中平稳时间序列法、马尔科夫链法和 BP 神经网络法预测结果历史符合率较高。本文的 Bayes 判别法的问题是如何筛选出与预报量关系密切的因子,这样其预报结果才可能与实况相符。

参考文献

- [1] 侯陶谦. 中国松毛虫[M]. 北京:科学出版社,1987:188-191.
- [2] 邹运鼎,程扶玖,查光济,等. 松针内含物与马尾松毛虫生存发育关系的研究[J]. 林业科学,1990,26(2):142-148.
- [3] 萧刚柔. 中国森林害虫[M]. 第 2 版. 北京:林业出版社,1992:948-953.
- [4] 张真,李典谟. 马尾松毛虫暴发机制分析[J]. 林业科学,2008,44(1):140-150.
- [5] 张爱兵,陈建,王正军,等. BP 网络模型和 LOGIT 模型在森林害虫预报上的应用初报[J]. 生态学报,2001,21(12):2159-2156.
- [6] PARK Y S, CEREHINO R, COMPIN A. Applications of artificial neural networks for patterning and predicting aquatic insect species richness in running waters [J]. Ecological Modelling, 2003, 160(3): 265-280.
- [7] 贾春生. 利用马尔可夫链方法测报马尾松毛虫发生级别[J]. 东北林业大学学报,2006,34(5):21-22.
- [8] ZHANG W J, ZHANG X Q, LIU G H. Recognizing spatial distribution patterns of grassland insects: neural network approaches [J]. Stochastic Environmental Research and Risk Assessment, 2008, 22(2): 207-216.

2731 - 2739.

[23] GRICE E A, KONG H H, CONLAN S, et al. Topographical and temporal diversity of the human skin microbiome [J]. Science, 2009, 324(5931): 1190 - 1192.

[24] SANCHEZ-CONTRERAS M, VLISIDOU I. The diversity of insect bacteria interactions and its applications for disease control [J]. Biotechnology and Genetic Engineering Reviews, 2008, 25(1): 303 - 344.

[25] BUCHNER P. Symbiosis in animals which suck plant juices, in endosymbiosis of animals with plant microorganisms [M]. New York: Interscience, 1965: 210 - 432.

[26] CHEN Deqiao, CAMPBELL B C, PURCELL A H. A new *Rickettsia* from a herbivorous insect, the pea aphid *Acyrtosiphon pisum* (Harris) [J]. Current Microbiology, 1996, 33 (2): 123 - 128.

[27] LI Tong, XIAO Jinhua, XU Zhaohuan, et al. A possibly new

Rickettsia-like genus symbiont is found in chinese wheat pest aphid, *Sitobion miscanthi* (Hemiptera: Aphididae) [J]. Journal of Invertebrate Pathology, 2011, 106(3): 418 - 421.

[28] FUKATSU T, TSUCHIDA T, NIKOH N, et al. *Spiroplasma* symbiont of the pea aphid, *Acyrtosiphon pisum* (Insecta: Homoptera) [J]. Applied and Environmental Microbiology, 2001, 67(3): 1284 - 1291.

[29] KOGA R, TSUCHIDA T, FUKATSU T. Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont *Buchnera* in an aphid [J]. Proceedings of the Royal Society B: Biological Sciences, 2003, 270(1533): 2543 - 2550.

[30] LI Tong, XIAO Jinhua, XU Zhaohuan, et al. Cellular tropism, population dynamics, host range and taxonomic status of an aphid secondary symbiont, SMLS (*Sitobion miscanthi* L type symbiont) [J/OL]. PLoS ONE, 2011, 6(7): e21944.

(责任编辑: 田 喆)

(上接 116 页)

[7] 谢兰芬,何鹏飞,吴毅歆,等. 解淀粉芽孢杆菌 B9601-Y2 对玉米叶斑病的防治效果[J]. 玉米科学, 2017, 25(2): 130 - 135.

[8] 崔文艳,何朋杰,尚娟,等. 解淀粉芽孢杆菌 B9601-Y2 对玉米的防病促生长效果研究[J]. 玉米科学, 2015, 23(5): 153 - 158.

[9] GARGE S S, NERURKAR A S. Evaluation of quorum quenching *Bacillus* spp. for their biocontrol traits against *Pectobacterium carotovorum* subsp. *carotovorum* causing soft rot [J]. Biocatalysis and Agricultural Biotechnology, 2017, 9: 48 - 57.

[10] LIU Chunming, YANG Zhenfu, HE Pengfei, et al. Deciphering the bacterial and fungal communities in clubroot-affected cabbage rhizosphere treated with *Bacillus subtilis* XF-1 [J]. Agriculture, Ecosystems & Environment, 2018, 256: 12 - 22.

[11] SHAO Jiahui, XU Zhihui, ZHANG Nan, et al. Contribution of indole-3-acetic acid in the plant growth promotion by the rhizospheric strain *Bacillus amyloliquefaciens* SQR9 [J]. Biology and Fertility of Soils, 2015, 51(3): 321 - 330.

[12] 何鹏飞. B9601-Y2 菌株的基因组解析及部分功能验证[D].

武汉: 华中农业大学, 2014.

[13] 汪汉成, 杨双剑, 许冬青, 等. 苯并噻二唑对漂浮育苗烟草白粉病的诱导抗性及其对烟苗安全性研究[J]. 植物保护, 2012, 38(6): 123 - 126.

[14] 邱德文. 生物农药研究进展与未来展望[J]. 植物保护, 2013, 39(5): 81 - 89.

[15] WANG Li, LI Xiaobao, SUO Haicui, et al. Soft rot of potatoes caused by *Bacillus amyloliquefaciens* in Guangdong province, China [J]. Canadian Journal of Plant Pathology, 2017, 39(4): 533 - 539.

[16] DES ESSARTS Y R, CIGNA J, QUÉTU-LAURENT A, et al. Biocontrol of the potato blackleg and soft rot diseases caused by *Dickeya dianthicola* [J]. Applied and Environmental Microbiology, 2016, 82(1): 268 - 278.

[17] HE Pengjie, CUI Wenyan, MUNIR S, et al. *Plasmodiophora brassicae* root hair interaction and control by *Bacillus subtilis* XF-1 in Chinese cabbage [J]. Biological Control, 2019, 128: 56 - 63.

(责任编辑: 田 喆)

(上接 128 页)

[9] 陈绘画, 王坚姬, 徐志宏, 等. 基于响应面方法的马尾松毛虫发生量混沌特性检测及其预测[J]. 东北林业大学学报, 2011, 39(9): 94 - 96.

[10] 田万银, 徐华潮. 浙江沿海防护林马尾松毛虫的预测预报模型[J]. 环境昆虫学报, 2012, 34(4): 401 - 406.

[11] 费海泽, 王鸿斌, 孔祥波, 等. 马尾松毛虫发生相关气象因子筛选及预测[J]. 东北林业大学学报, 2014, 41(1): 136 - 140.

[12] 王庆, 毕猛, 杜婷, 等. 基于气象因子的马尾松毛虫发生率空间格局研究[J]. 林业科学研究, 2016, 29(2): 256 - 260.

[13] 张孝羲. 昆虫生态及预测预报[M]. 第 3 版. 北京: 中国农业出版社, 2006: 205 - 221.

[14] 许章华, 李聪慧, 刘健. 马尾松毛虫害等级的 Fisher 判别分析[J]. 农业机械学报, 2014, 45(6): 401 - 406.

[15] 徐克学. 生物数学[M]. 北京: 科学出版社, 2002: 40 - 50.

[16] 周夏芝, 王振兴, 余燕, 等. 马尾松毛虫幼虫高峰期发生量的预测模型研究[J]. 应用昆虫学报, 2017, 54(6): 1031 - 1043.

(责任编辑: 田 喆)